

System and Method for Direct Multi-Modal Annotation of Objects

Abstract of the Disclosure

[0076] The system includes an image display system, a direct annotation creation module, an annotation display module, a vocabulary comparison module and a dynamic updating module. These modules are coupled together by a bus and provide for the direct multi-modal annotation of media objects. The direct annotation creation module creates annotations in response objects. The image display system is coupled to a media object cache and displays images to user input and stores the annotations in memory. The annotation display module works in cooperation with the image display system to display the annotations themselves or graphic representations of the annotation positioned relative to the images of the objects. The vocabulary comparison module works in cooperation with the direct annotation creation module to receive audio input and present matches of annotations. Similarly, the dynamic updating module receives input annotations, and updates an audio vocabulary to include a text annotation for new audio input signal. The system provides direct annotation of images. Once an image is displayed, the user need only select an image and speak to create an annotation. The system automatically creates the annotation, associates it with the selected images, and displays either a graphic representation of the annotation or a text translation of the audio input. The present invention may also present likely matches of text to the audio input and/or update an audio vocabulary in response to input of audio inputs that are not recognized.